The innovations in the realm of Artificial Intelligence (AI) and Machine Learning (ML) and its application in real-world comes with their own challenges. For instance, a research study on scaling laws for neural language models shows that the increasing complexity and scale of ML models, coupled with the exponential growth of data, have created a pressing need for more powerful hardware and software solutions. This calls for developing secure and efficient distributed training algorithms, infrastructure and frameworks for utilizing large-scale computing resources for training complex models to maximize performance and minimize latency, energy and computational cost. Many challenges need to be addressed in the next few years, and I want to pursue a PhD program to solve these challenges.

With a strong interest in hardware and systems, my undergraduate program in Mechanical Engineering afforded me the chance to garner a plethora of technical skills through courses in C and C++ Programming, Numerical Methods, Computer Graphics and Advanced Mechatronics. Furthermore, I completed courses in Data Structures and Algorithm, Advanced Calculus, Differential Equations, Probability and Statistics, Electronics and Electrical Engineering, Electrical Machines and Advanced Linear Algebra among other computing and applied courses. With Honours, I graduated amongst the top percentile of students in the department.

Inspired by leading researchers in nonlinear dynamic systems, I collaborated with my undergraduate research professor, Dr. Rabindra Kumar Behera, to investigate the coupled nonlinear behavior of a beam structure with a moving mass. I conducted a literature study around dynamic analysis of beams, to understand the current state of research and identify gaps that my study could address. I then outlined my objective to analyze the effects of different boundary conditions on the dynamic behavior of beams with moving masses and investigate specific nonlinear phenomena. The challenge to develop accurate mathematical models that capture the nonlinear dynamics of the system was convoluted. Based on the objective and literature study, first, I developed a mathematical model based on the Euler-Bernoulli beam assumption and then applied Hamilton's principles that led to an analytical nonlinear geometric model incorporating the motion of the beam with a moving mass. A Galerkin discretization method was applied to convert the continuous equations into a discrete form for numerical analysis which was a crucial step. It was highly challenging to find methods that can achieve stability and convergence on very nonlinear systems. With further literature study on implicit methods and sensitivity analysis, a numerical method, perturbation method of multiple scales, a method highly known for stability in nonlinear systems, was applied to solve the resulting ordinary differential equations to analyze the system over time. With the ODE software optimization in MATLAB I simulated the system to visualize the variation of amplitude, deflection, and other parameters with respect to time and analyzed the resonance and nonlinear dynamic interactions on boundary conditions. Optimization of the program, vectorization, symbolic computation and invoking batch processing helped to address the challenge of computations that required significant resources. This was followed by benchmarking with existing analytical and numerical solutions from the literature to ensure the accuracy and reliability of the results. This research with its potential applications in engineering design and small-scale electronic devices, culminating in a co-authored publication, solidified my interest in data analysis and algorithm development. My involvement in this acoustics research project and by witnessing the power of computational tools to solve real-world problems through my first-hand experience with software programming for solving highly complex and computationally intensive models, ignited my passion to learn how a hardware system with software optimizations efficiently performs computational tasks.

This prompted me to embark upon earning a master's degree in Mechanical Engineering with Data Science specialization at the University of Washington. I had the opportunity to develop strong foundation in AI and ML through graduate research courses on Computer Vision, Natural Language Processing (NLP), Software Development for Data Scientists, Deep Learning, Applied Parallel Computing along with a course on Database Management Systems and Linear Systems Theory. I also gained valuable research experience working on course projects on visual dynamics and physical reasoning by representation learning, aspect based sentiment analysis with transformer encoder-decoder and BERT model, vision analysis on lane detection model using ResNet model, deep reinforcement learning for developing trading algorithm and ML informed analysis on material science. My project on optimizing a stereoscopic depth perception algorithm using parallel computing techniques was one of the notable ones. I started with the sequential implementation of disparity calculation to understand the algorithm followed by improving the sequential performance of the algorithm by gradient computation. This improved the accuracy of the model but at the cost of higher computational time. We then examined the algorithm and identified computationally intensive blocks to parallelize it. There were several challenges, for instance: memory bandwidth limitations when transferring large image data, GPU architecture constraints requiring careful thread block sizing, complex synchronization between threads when computing gradients, memory access pattern optimization, handling edge cases in parallel implementation. We experimented and applied advanced CUDA features with coalesced memory access in CUDA kernels, algorithm optimization, boundary checking and optimal thread block dimensions analysis to achieve higher performance, efficient resource utilization and thread mapping to image pixels. Additionally, I applied CUDA graphs on repetitive kernel launches and memory operations where with a single CPU operation the work was pre-scheduled on GPU as a graph network to avoid multiple kernel calls. This eliminated launch latency for the disparity calculation kernels and reduced CPU-GPU synchronization overhead. This resulted in an additional 8% speedup on average and further validated the results with the baseline sequential algorithms.

To enhance my academic foundation, I gained practical experience as a Machine Learning Intern at Chubb Insurance, contributing to projects involving language inference, entity detection, knowledge graph construction, information extraction, document classification, and summarization. A significant project was the development of an anomaly detection model using a probabilistic unsupervised learning approach. I conducted a comprehensive literature review on a PyOD paper, a benchmarking study of various anomaly detection models having linear, proximity-based, ensembles, neural nets and graph-based models. Based on the study, I selected an empirical cumulative distribution function based outlier detection model through the detailed study of its paper. This model having unsupervised learning architecture excelled in handling large, varying distribution datasets, offering scalability, maintainability and higher efficiency. It achieved an accuracy of 83% in detecting anomalies in a dataset of over half a million data points. One another notable opportunity was when I led a team of diverse developers on the Database Development Project for the Institute for Health Metrics and Evaluation (IHME). This unique industry experience exposed me to learn about best practices in database management, high-performance data storage, and ML pipeline optimization and at the same time I gained valuable leadership and team management skills, fostering collaboration and inclusivity among team members. I also developed technical expertise in ER diagrams, data analysis, database performance optimization, and data integrity. This project ignited my passion for database systems, leading me to pursue a database management course and build an interactive UI web application with a SQL database on Azure. Through this experience, I honed my skills in database design, query optimization, and collaborative problem-solving. Beyond technical skills, I am committed to social impact through promoting shared values, inclusivity, diversity and accessibility in engineering, as demonstrated by my work with a non-profit organization in the database administration and IT team.

My academic journey, beginning with a Mechanical Engineering undergraduate degree with research and culminating in a Master's with specialization in Data Science and Machine Learning, has cultivated a deep interest at the intersection of hardware and software domains. I am passionate about exploring the synergies between these domains and am drawn to the research at the forefront of hardware-software co-design, distributed systems, and machine learning. I am particularly persuaded to the challenge of optimizing large language-vision (LLVM) models for resource-efficient and resource-constrained systems by leveraging insights from both theoretical and empirical analysis of learning algorithms and foundational models. To this end, I aim to explore techniques on how we can integrate hardware acceleration and parallelization techniques with software optimization methods for the computational kernels of LLVM models to achieve maximum throughput and performance gains with efficient memory utilization. Additionally, to address the challenge of finding optimal architectural choices and specialized hardware accelerator design for balancing performance, hardware heterogeneity, energy efficiency, communication and synchronization overhead optimization, resource utilization and cost-effectiveness when designing or applying to scalable and fault-tolerant distributed ML. Furthermore, another avenue of exploration within this research that I am interested in is the challenge of ensuring robust security and privacy in highly distributed and parallel ML systems, while federated learning and adversarial training has gained significant attention. A PhD program in Computer Science would enable me to delve deeper into this field, pioneer original research, and contribute to its growth. This will help me transition into a successful research career in Computer Science, in academia or industry, by providing me with the necessary skills, knowledge, and professional connections. I am interested in these areas because I can contribute to the wider adoption of AI in various industries, from healthcare, finance and environmental sustainability to autonomous systems and technology and push the boundaries of ML systems, leading to novel architectures, algorithms, and techniques that outperform existing solutions. Additionally, this integrated and interdisciplinary perspective allows me to explore innovative solutions that bridge the gap between hardware and software, as well as between machine learning and security that would enable me to tap into lot more established academic research communities and industry for support and learning. In the long-term I want to utilize my PhD to weave a career in AI/ML research mainly as a researcher or scientist in industry or in academia to leverage my interest in open-source projects, writing research papers, building new technologies where I can contribute new knowledge and conduct original research to solve complex AI challenges that can have significant impact in various industries.

Through collaborative interactions with advisors and peers, which I now see as a critical foundation for research, I have developed a strong foundation in research methodology and a keen eye for identifying and addressing research gaps. My ability to recognize and leverage subtle insights with adapting to evolving research challenges, coupled with persistence in overcoming challenges and embracing diverse perspectives, has led to significant breakthroughs in my research. Having lived in three countries encompassing global and unique perspective on the possibilities of how technology can change lives, the institute with its strong research community and leading experts, would be an ideal environment to mold my intellectual ability, challenge me to go beyond and assist me to accomplish my aspirations as a researcher. I believe my diverse background and interdisciplinary skill set with broad spectrum of knowledge in engineering and data science will enable me to contribute meaningfully to your department's ongoing cutting-edge AI research and bring a unique dimension to your community.